UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/761,222	01/16/2001	Seiichiro Abe	1990.65128	4752

24978 7590 01/14/2004

GREER, BURNS & CRAIN
300 S WACKER DR
25TH FLOOR
CHICAGO, IL 60606

EXAMINER	
ALAM, SHAHID AL	

ART UNIT	PAPER NUMBER
2172	8

DATE MAILED: 01/14/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/761,222	ABE, SEIICHIRO	
	Examiner	Art Unit	
	Shahid Al Alam	2172	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *17 October 2003*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-9* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-9* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. §§ 119 and 120

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

13)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application) since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

a) ☐ The translation of the foreign language provisional application has been received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121 since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

Response to Arguments

1. Applicant's arguments filed October 17, 2003 have been fully considered but they are not persuasive for the following reasons.

2. Applicant argue that WebMate does not disclose a search condition designating unit that designates a file as a search condition and transmits contents of a designating file via a network for a search requesting source in combination with a document search unit that forms a keyword from the file contents transmitted from a search conditioned designating unit that searches similar documents.

WebMate searches documents based on a similarity between a user profile and a document, and not according to a similarity between a document and another document.

Examiner respectfully disagrees the entire allegation as argued. Examiner, in his previous office action, gave detail explanation of claimed limitation and pointed out exact locations in the cited prior art.

WebMate teaches an agent that helps users to effectively browse and search the Web.

WebMate extends the state of the art in Web-based information retrieval in many ways. It uses multiple TF-IDF vectors to keep track of user interests in different domains. These domains are automatically learned by WebMate. WebMate uses the Trigger Pair-Model to automatically extract keywords for refining document search.

During search, the user can provide multiple pages as similarity/relevance guidance for the search. The system extracts and combines relevant keywords from these relevant pages and uses them for keyword refinement. Using these techniques, WebMate provides effective browsing and searching help and also compiles and sends to users personal newspaper by automatically spiding news sources. WebMate utilizes TF-IDF method with multiple vectors representation. The basic idea of the algorithm is to represent each document as a vector in a vector space so that documents with similar content have similar vectors. Each dimension of the vector space represents a word and its weight. The values of the vector elements for a document are calculated as a combination of the statistics term frequency *TF(w, d)* (the number of times word w occurs in document d) and document frequency *DF(w)* (the number of documents the word w occurs in at least once). From the document frequency the inverse document frequency *IDF(w)* can be calculated.

One of the most important ways in which current information retrieval technology supports refining searches is relevance feedback. Relevance feedback is a process where users identify relevant documents in an initial list of retrieved documents, and the system then creates a new query based on sample relevant documents. The idea is that since the newly formed query is based on documents that are similar to the desired relevant documents, the returned documents will indeed be similar. The central problems in relevance feedback are selecting "features" (words, phrases) iZ1 from relevant documents and calculating weights for these features in the context of a new query. In WebMate agent, the context of the search keywords in the "relevant" web

pages is used to refine the search because a user tells the system some page is relevant to his search, the context of the search keywords is more informative than the content of the page.

For the above reasons, Examiner believed that rejection of the last Office action was proper.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 1-2, 4-5, and 9 are rejected under 35 U.S.C. 102(b) as being anticipated by the publication, "WebMate: A Personal Agent for browsing and Searching," Chen et al., Proceedings of the 2nd International Conference on Autonomous Agents, May, 1998, NY, USA, ACM Press, pages 132-139, hereinafter "WebMate."

With respect to claim 1, WebMate teaches a document information search apparatus for searching document information on the basis of a search request transmitted through a network (Page 134, Fig. 1; search requests are made in the WWW) and responding, wherein: a search condition designating unit which designates a file as a search condition (page 134, col. 2, lines 4; page 137, col. 2, lines 13-18 ("the

context of the search keywords in the relevant web pages is used"); a user designates a URL); and transmits contents of said designated file via the network is provided for a search requesting source (page 134; Fig. 1; WebMate receives the web page designated by a user); and a document search unit which forms a keyword from the file contents transmitted from said search condition designating unit (page 134, col. 2, the 2nd paragraph; WebMate constructs a query based on a current profile which is formed of the keywords that come from a plurality of domains including the Web page visited by users when the users designate them; the creation of a personal profile is described in page 133, col. 12, section 3.1) and
searches similar documents from a database (page 134, col. 2; WebMate calculates similarity between the profile and a plurality of Web pages, and recommends the ones based on a threshold; note that WebMate searches a plurality of URL's of users do not designate any particular Web page or URL) is provided on a search side.

As to claim 2, WebMate teaches an apparatus according to claim 1, wherein said search condition designating unit transmits a head file portion of the designated file contents (page 137, lines 13-18; since the designated file is a Web page, the URL associated with the designated Web page is considered the head file)
Claim 4 recites the following:

an apparatus according to claim 1, wherein index information describing a list of important words extracted from search target documents is stored for every document in said database, and said document search unit on the search side comprises: a text extraction processing unit which extracts a text document from the file contents received

in response to the search request; a morpheme analyzing unit which extracts nouns by a morpheme analysis of said text document; a keyword forming unit which extracts important words from said nouns and forms a keyword in which said important words are coupled by OR; and a search executing unit which searches similar documents by searching the search database by said keyword and notifies the search requesting source of a search result.

With respect to the limitations of claim 4, WebMate teaches creation of profiles and generation of relevant Web pages by extracting keywords from the relevant Web pages by using TF-IDF (term frequency -inverse document frequency) method. The TF-IDF requires that all documents be parsed for extracting keywords including nouns, and excluding the stop words, and also requires that documents be ranked in a particular order. As to the step of notifying the search, see page 137, col. 2, and page 138, col. 1, wherein a list of 5 relevant documents is provided.

As to claim 5 (an apparatus according to claim 4, wherein said keyword forming unit counts the number of times of appearance showing in which documents in the index of each of the search documents stored in said document database each of said nouns appears, selects a predetermined number of upper words each having the number of times of appearance in a predetermined range, and forms the keyword), WebMate teaches the use of TFIIDF method and in addition, teaches the use of "top 5 words" in documents for retrieval of the most relevant documents (page 138, col. 1, line 15-17).

As to claim 9 (an apparatus according to claim 1, wherein said search condition designating unit of said search requesting source is provided by a WWW browser of a client, transmits the contents of the file designated by a search request picture plane of said WWW browser to a search machine of a WWW server through the network, and sends said file contents to said document search unit), WebMate shows the WWW environment in page 134, Figure 1. In accordance with the description provided on page 7, lines 18-22 of the Applicant's Disclosure, it appears that the "search request picture plane" is nothing more than a query box where a keyword can be typed in by a user. Since WebMate teaches a browser, it inherently teaches the query box and/or search request picture plane as claimed.

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over WebMate.

Claim 6 (an apparatus according to claim 5, wherein in the case where the number of documents in the index is assumed to be (N), said keyword forming unit selects upper ten words each having the number (H) of times of appearance in a range where 2N/3.gtoreq.H.gtoreq.1 and forms the keyword) requires that top 10 keywords be

used to rank and present most relevant documents. WebMate suggests that top 5 keywords be used (page 138, column 1, line 15 – 17).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to select top 10 instead of top 5 because such a change can be adopted without reconfiguring the WebMate system or without incurring any reconfiguration overhead, while a person of ordinary skill in the art would find this as an added flexibility of the system.

6. Claims 7 and 8 are rejected under 35 U.S.C. 103(a) as being unpatentable over WebMate as applied to claims 1, 4 and 5 above, and further in view of the publication, "CiteSeer: An Autonomous Web Agent for Automatic Retrieval and identification of Interesting Publications," by Bollacker et al., proceedings of the International Conference on Autonomous Agents, May 1998, ACM Press, pages 116-123, hereinafter "Citeseer."

As to claim 7 (an apparatus according to claim 5, wherein said keyword forming unit allows property information extracted from the file received in response to the search request to be included in said keyword, thereby allowing the similar documents to be searched), WebMate discloses the extraction of keywords (WebMate teaches that a user can provide any URLs that he would like to be the information sources and that the chosen URL may be used to expand the search in page 134, col. 2, lines 4-6), but does not explicitly indicate that the keywords include property information as claimed.

As to claim 8 (an apparatus according to claim 7, wherein said property information includes a writer of the file received in response to the search request, and

a document title.), WebMate discloses the extraction of keywords, but does not explicitly indicate that the keywords extracted include the writer of the file or the title of the file.

As to claims 7 and 8, Citeseer uses a sub-agent to search a plurality of Web pages when a broad keyword is entered by a user in the search query (page 118, col. 1, section 3.13, lines 5-6). Citeseer further teaches the extraction of title and author in response to submitted query (page 118, col. 2, the bottom paragraph).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine WebMate and Citeseer to make the system user friendlier as such the user will be able to see the bibliographic information of to-be-retrieved documents. It would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine WebMate and Citeseer to eliminate some of the retrieval candidates (i.e., to-be-retrieved documents) to avoid the down loaning and transmission overhead. It is general knowledge available to one of ordinary skill in the data processing art that a document is more likely to have bibliographic information and that a retrieval system incurs overhead to download a document.

7. Claim 3 is rejected under 35 U.S.C. 103(a) as being unpatentable over WebMate as applied to claim 1 above, and further in view of U. S. Patent No. 6,182,085 issued to Eichstaedt ("the `085 patent").

As to claim 3 (an apparatus according to claim 1, wherein said search condition designating unit allows an HTML file and an Excel file to be included in the file which is designated as said search condition), WebMate teaches the parsing of an HTML page (page 134, col. 1, section 3.2. lines 6), however does not explicitly indicate the

processing of an EXCEL file.

As to the limitation, "...said search condition designating unit allows an HTML file and an Excel file to be included in the file which is designated as said search condition", WebMate does not explicitly indicate that it is capable of parsing a EXCEL file submitted a query. With respect to claim 3, the `085 patent (Eichstaedt et al.), in column 5, lines 12-32, teaches:

One example of a Gatherer 302 communicatively linked to a web 304 is pictured in FIG. 3 and has a number of components. The web 304 may comprise an Internet, an intranet, or a single information source including media or multimedia objects. The Gatherer 302 may include a Crawler 306 component that crawls media sources and retrieves objects while a Recognizer 308 component tries to determine the format for each of the retrieved objects. A Summarizer 310 component **contains specialized codes that enable it to read a great number of different object formats such as a Freelance graphics presentation, an HTML page, a Lotus Notes database, or an Excel spreadsheet.** It also provides a flexible structure for plugging -_n customized summarization codes to be used for summarizing data from a specific :Location. Compressed files included in a ZIP, TAR or JAR file are first extracted out by an Expander 312 component and then processed by the Summarizer 310. A Gatherer may also carry an embedded HTTP server (not shown) so that system administrators can use a web-browser to control its operations and monitor its status.

Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to combine WebMate and the '085 patent to make the system more user-friendly as such the user would be to submit queries in any form and would not have to convert the query object into a specific format to search for relevant information.

Conclusion

8. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Contact Information

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Shahid Al Alam whose telephone number is (703) 305-2358. The examiner can normally be reached on Monday-Thursday 8:00 A.M. - 4:00 P.M..

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John E Breene can be reached on (703) 305-9790. The fax phone number for the organization where this application or proceeding is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-3900.

Shahid Al Alam
Primary Examiner
Art Unit 2172

11 January 2004